actuators

MDPI

Article

Physics-Constrained Dual-Attention Reinforcement Learning for Semi-Active Lateral Vibration Control of High-Speed Trains

Dongyu Fan, Lei Gao, Runliang Tian *, Yiwei Zhao and Zhaoyang Xing

School of Mechanical Engineering, Shijiazhuang Tiedao University, Shijiazhuang 050043, China
* Correspondence: tianrunliang@stdu.edu.cn

Abstract

High-speed trains are prone to severe lateral vibrations induced by track irregularity excitations under complex operating conditions, which deteriorate ride comfort, reduce running stability, and accelerate wheel–rail wear. Existing lateral suspension vibration control methods mainly rely on fixed parameters and expert experience. To further optimize and improve the control performance of semi-active lateral suspension systems for high-speed trains, this paper proposes the Physics-Constrained Dual-Attention Reinforcement Learning for Semi-Active Lateral Vibration Control of High-Speed Trains (SA-PRL). The proposed method is built upon a semi-active Twin Delayed Deep Deterministic Policy Gradient algorithm (SATD3). To ensure the physical realizability of control actions, the Logical Constraints of Skyhook Control (LCSC) are introduced to map the controller output into physically feasible damping commands. Furthermore, to enhance the perception of critical state features and improve value estimation capability, a Critic Network with Dual-Head Self-Attention (CDHA) is developed. Based on a railway vehicle lateral dynamic model, a series of simulation experiments are conducted under four excitation conditions, including single-peak sinusoidal excitation and multi-peak sinusoidal excitation, as well as the Chinese high-speed railway track irregularity (CHSRTI) and German low-interference track irregularity (GLITI) excitations. The results demonstrate the lateral vibration suppression capability of the proposed method over passive suspension, skyhook damping control, and existing reinforcement learning methods, with reductions of 47.5% and 47.7% in the RMS carbody lateral acceleration under the CHSRTI and GLITI spectra, respectively. The proposed method provides an effective solution for semi-active lateral vibration control of high-speed trains.

Keywords: high-speed train; lateral vibration; deep reinforcement learning; semi-active suspension; track irregularity

Check for updates

Academic Editors: Tongfei Tian, Yingdan Liu, Shou-Hu Xuan and Yuhiro Iwamoto

Received: 3 July 2026
Revised: 31 July 2026
Accepted: 10 August 2026
Published: 12 August 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

With the continuous increase in operating speeds, high-speed rail transportation has become a major direction for the development of modern railway systems. However, higher speeds amplify the effects of track irregularities and complex external disturbances on vehicle dynamics, making the carbody more susceptible to severe lateral vibrations. These vibrations not only deteriorate ride comfort but also accelerate wheel–rail wear and structural fatigue damage [1,2]. Therefore, effective suppression of lateral vibrations is essential for ensuring the safety, stability, and comfort of high-speed railway systems, making it a critical research topic in the field of railway vehicle dynamics [3–5]. As shown

in Figure 1, the research background of high-speed train lateral vibration and the overall framework of the proposed semi-active lateral vibration control method are illustrated.

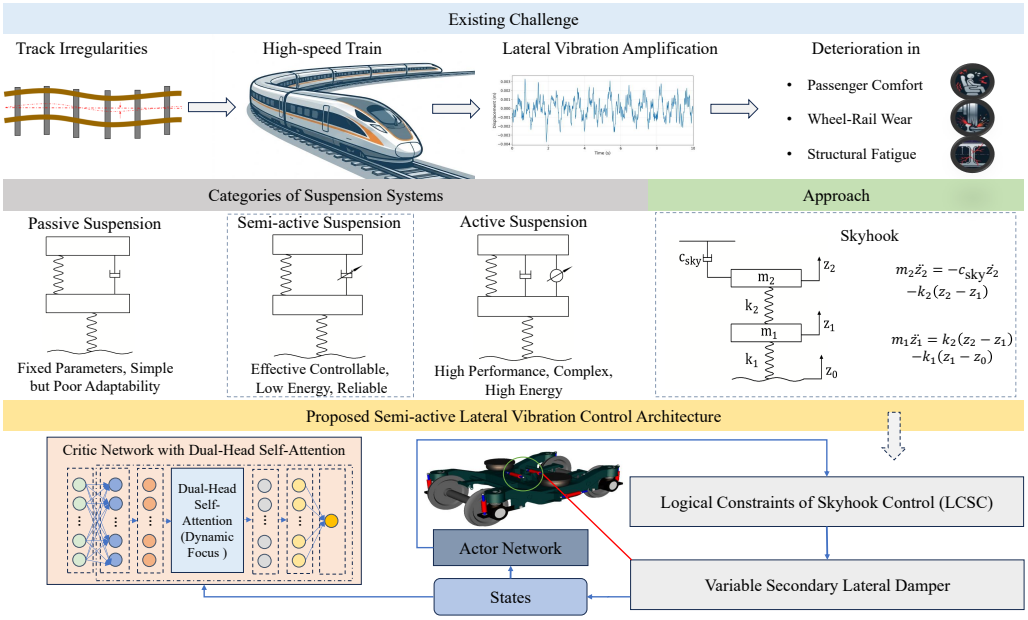


Figure 1. Research motivation and overview of the proposed physics-constrained deep reinforcement learning method for semi-active lateral vibration control of high-speed trains. Broad solid-line arrows describe how track irregularities excite train lateral vibration and its negative effects. Blue solid arrows denote neural-network data flow, and the dashed arrow transmits skyhook-derived constraints. The ellipsis inside the critic network omits hidden-layer neurons.

To effectively suppress lateral vibrations in high-speed trains, suspension systems play a crucial role in vibration isolation, impact mitigation, and vehicle stability enhancement [6]. According to their control mechanisms, railway suspension systems can be broadly classified into passive, active [7,8], and semi-active suspensions [9,10]. Passive suspensions rely on fixed spring and damping parameters to attenuate vibrations, offering a simple structure but limited adaptability to varying operating conditions. Active suspensions can generate control forces directly and generally provide vibration reduction performance; however, they are characterized by increased system complexity and energy consumption. In contrast, semi-active suspensions regulate damping characteristics in real time, achieving a favorable balance between control effectiveness, energy efficiency, and system reliability. Consequently, they are widely regarded as a practical and effective solution that combines satisfactory vibration control performance with engineering feasibility [11]. Extensive research efforts have been devoted to improving the control performance of semi-active lateral suspensions for high-speed trains. Among the various control strategies, skyhook-based methods have become one of the most widely adopted approaches owing to their clear physical interpretation and reliable vibration attenuation capability [12,13]. To overcome the limitations of fixed damping parameters in existing skyhook control, numerous studies have integrated advanced control theories to dynamically optimize suspension damping forces, thereby further enhancing lateral stability and ride comfort under diverse operating conditions [14,15].

Although existing semi-active lateral suspension control methods have achieved significant improvements in vibration attenuation for high-speed trains, most approaches still rely on accurate dynamic models, fixed control parameters, and expert-designed control rules. However, under complex operating conditions, vehicle dynamics exhibit pronounced nonlinear and time-varying characteristics due to track irregularities and speed fluctuations.

Consequently, existing control strategies often struggle to adapt effectively to changing environments, limiting further improvements in control performance and robustness. Therefore, there is a growing need for intelligent control methods with autonomous learning and online decision-making capabilities [16] to enhance the adaptability of high-speed train lateral vibration control systems. Recent advances in artificial intelligence have promoted the development of Deep Reinforcement Learning (DRL) [17–19], which has demonstrated considerable potential for complex nonlinear control problems owing to its model-free nature, adaptive decision-making capability, and end-to-end optimization framework [20–22]. Representative DRL algorithms, including Deep Deterministic Policy Gradient (DDPG) [23,24], Twin Delayed Deep Deterministic Policy Gradient (TD3) [25–27], and their improved variants [28–30], provided new avenues for intelligent suppression of lateral vibrations in high-speed trains. Nevertheless, existing DRL-based suspension control studies have primarily focused on automotive systems [31–33]. Compared with automotive suspensions [34–36], high-speed train lateral dynamics involve stronger multi-degree-of-freedom coupling effects, more complex track-induced excitations, and more stringent safety requirements [37]. As a result, the application of DRL to high-speed train lateral semi-active suspension control remains largely unexplored, and a mature and effective intelligent control framework has yet to be established.

Although DRL exhibits strong autonomous decision-making and online optimization capabilities [38], existing DRL-based control policies generally lack explicit physical constraints derived from established damping mechanisms. As a result, the learned control actions may violate the physical characteristics and operational limitations of suspension systems, thereby hindering practical engineering implementation. In contrast, skyhook damping control provides a clear physical interpretation and has demonstrated effective vibration attenuation performance [39]. Its damping-switching mechanism can serve as an effective physical constraint for DRL-based control policies, guiding the learning process toward physically feasible control actions while preserving the adaptive optimization capability of reinforcement learning. Therefore, integrating the skyhook damping mechanism with DRL offers a promising approach for semi-active suspension control of high-speed trains, enabling a balance between vibration control performance and engineering practicality.

Based on the above considerations, this study proposes the Physics-Constrained Dual-Attention Reinforcement Learning for Semi-Active Lateral Vibration Control of High-Speed Trains (SA-PRL). A critic network incorporating a dual-head attention mechanism is developed to improve the perception and evaluation of complex dynamic state features. In addition, the skyhook damping mechanism and semi-active suspension physical constraints are integrated into the reinforcement learning framework, where the operational boundaries of the variable damper are imposed at the execution layer as terminal physical constraints.

The main contributions of this study are summarized as follows:

- A physics-constrained deep reinforcement learning method designed for the characteristics of high-speed train lateral vibrations is developed to improve adaptive decision-making and vibration suppression performance under complex operating conditions.
- By combining the logical constraint of skyhook damping control with the continuous optimization capability of reinforcement learning, an adaptive damping output function is designed to achieve coordinated vibration control under physical constraints.
- A critic network with dual-head self-attention is developed to enhance the perception and representation of complex dynamic state features, thereby improving control decision-making capability.

2. Vehicle Model and Track Irregularity Excitation

Research on reinforcement learning suspension control involves different modeling strategies to balance computational efficiency and dynamic fidelity. The complete railway vehicle models with detailed wheel–rail coupling and actuator nonlinearities deliver comprehensive engineering evaluation capacity for railway suspension controllers [40]. Meanwhile, model order reduction techniques demonstrate that low-dimensional simplified models can drastically cut iterative training costs while retaining core vibration characteristics, making them suitable for preliminary algorithm exploration [41]. Modeling of train lateral vibration control under track irregularity excitation is a key research topic in railway vehicle dynamics and forms the foundation for intelligent control development. In this study, a two-degree-of-freedom quarter suspension model is employed to construct a train lateral dynamic simulation environment, which provides a reliable basis for the training and validation of subsequent deep reinforcement learning algorithms.

2.1. Lateral Dynamics Modeling of Trains

The lateral dynamics modeling of trains is fundamental for analyzing lateral vibration characteristics and developing vibration control strategies. Based on the quarter suspension model, the train lateral dynamic model is established, as shown in Figure 2, where m_c and m_b denote the mass of the carbody and the bogie respectively, k_p represents the primary lateral stiffness, k_s and c_s denote the secondary lateral stiffness and the secondary lateral damping respectively, x_c and x_b are the lateral displacement of the carbody and the lateral displacement of the bogie, respectively, x_w represents the lateral displacement induced by track irregularities, and f_a denotes the lateral control force generated by the semi-active suspension through the modulation of the variable damper c_s. The specific parameters are taken from a certain high-speed train, as listed in Table 1 [42].

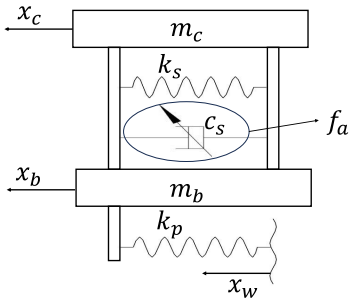


Figure 2. A quarter railway vehicle lateral model.

Table 1. Parameters of the quarter of the suspension model.

Symbol	Description	Value
m_c	Mass of carbody	8411.5 kg
m_b	Mass of bogie frame	1140 kg
k_p	Primary lateral stiffness	1.96×10^6 N/m
k_s	Secondary Lateral Stiffness	1.66×10^5 N/m
c_s	Secondary Lateral Damping	3.0×10^4 N · s/m

Based on Newton's second law, the lateral dynamic equations of the train can be formulated as follows:

$$m_c\ddot{x}_c = -k_s(x_c - x_b) + f_a \quad (1)$$

$$m_b\ddot{x}_b = k_s(x_c - x_b) - k_p(x_b - x_w) - f_a \quad (2)$$

During the control process, the lateral control force f_a generated by the semi-active suspension through modulation of the variable damper c_s satisfies the following relationship:

$$f_a = c_s(\dot{x}_c - \dot{x}_b) \tag{3}$$

The state-space representation of the suspension system can be expressed as:

$$\begin{cases} \dot{x} = Ax + BU \\ y = Cx + DU \end{cases} \tag{4}$$

The four-dimensional state vector is constructed using the lateral displacement and lateral velocity of the carbody, as well as the lateral displacement and lateral velocity of the bogie.

$$x = \begin{bmatrix} x_c \\ \dot{x}_c \\ x_b \\ \dot{x}_b \end{bmatrix} \quad y = \begin{bmatrix} x_c - x_b \\ \ddot{x}_c \end{bmatrix} \quad U = \begin{bmatrix} x_w \\ f_a \end{bmatrix}$$

Based on the dynamic equations, the state-space representation can be obtained as follows:

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\frac{k_s}{m_c} & 0 & \frac{k_s}{m_c} & 0 \\ 0 & 0 & 0 & 1 \\ \frac{k_s}{m_b} & 0 & -\frac{k_s+k_p}{m_b} & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{m_c} \\ 0 & 0 \\ \frac{k_p}{m_b} & -\frac{1}{m_b} \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 0 & -1 & 0 \\ -\frac{k_s}{m_c} & 0 & \frac{k_s}{m_c} & 0 \end{bmatrix} \quad D = \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{m_c} \end{bmatrix}$$

2.2. Track Irregularity Excitation

The power spectral density (PSD) of track irregularities is a fundamental tool for characterizing track excitation properties. Various countries have developed PSD models tailored to their own railway systems, including the Chinese high-speed railway track irregularity (CHSRTI) spectrum, the Federal Railroad Administration track irregularity spectrum of the United States, the German low-interference track irregularity (GLITI) spectrum, and the Japanese standard track irregularity spectrum. In this study, the lateral CHSRTI spectrum is modeled. Its one-sided PSD is expressed as [43]:

$$S(f) = \frac{A}{f^k} \tag{5}$$

where f denotes the spatial frequency (1/m), A is the track irregularity coefficient, and k is the spectral exponent of the track irregularity, as listed in Table 2.

Table 2. Parameters of the lateral CHSRTI spectrum.

Segment	A	k	Cut-Off Spatial Frequency (1/m)	Cut-Off Wavelength (m)
1	3.9513×10^{-3}	1.8670	0.0450	22.2
2	1.1047×10^{-2}	1.5354	0.1234	8.1
3	7.5633×10^{-4}	2.8171	—	—

The generation of time-domain samples of track irregularities is performed using the Inverse Fast Fourier Transform (IFFT) method. Based on the frequency-domain energy equivalence principle, this approach constructs a complex spectrum that satisfies the target

PSD characteristics and assigns random phases to efficiently simulate the time-series representation of track irregularities. First, the spatial PSD is converted into a temporal one-sided PSD under the energy conservation principle, with the following relationship:

$$G(f_t) = \frac{S(f)}{V} \tag{6}$$

where f_t denotes the temporal frequency (Hz), V is the train speed (m/s), and $G(f_t)$ is the transformed one-sided PSD in the time domain.

To satisfy the conjugate symmetry of real-valued signals, the one-sided spectrum is converted into a two-sided PSD, expressed as $S_x(f_t) = G(f_t)/2$. Based on the relationship between frequency-domain sampling intervals and energy preservation, the amplitude of the complex spectrum is constructed as

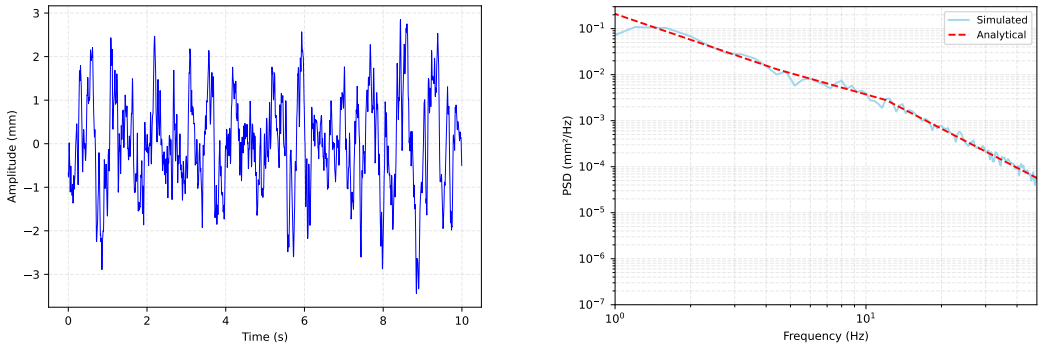
$$|X_k| = N_r \sqrt{S_x \cdot \Delta f} \tag{7}$$

where N_r denotes the total number of sampling points, and Δf is the frequency resolution. Uniformly distributed random phases in $[0, 2\pi]$ are then generated to construct a discrete complex spectrum that satisfies the target spectral characteristics. Finally, the IFFT is applied to obtain the time-domain track irregularity samples,

$$x(t) = \mathrm{IFFT}[X(k)] \tag{8}$$

where $\mathrm{IFFT}(\cdot)$ denotes the inverse fast Fourier transform operator.

To ensure simulation accuracy, appropriate frequency truncation limits are defined, where the lower cut-off frequency corresponds to a long-wavelength irregularity of 100 m, and the upper cut-off frequency corresponds to a short-wavelength irregularity of 2 m, thereby covering the vibration-sensitive frequency range of the train. The periodogram method is employed to estimate the PSD of the simulated signal, which is then compared with the theoretical analytical spectrum for validation. The simulation results, as shown in the Figure 3, indicate that the generated time-domain samples are in strong agreement with the target analytical spectrum. This demonstrates that this method can accurately reproduce the statistical characteristics of the CHSRTI spectrum and the results can serve as a reliable excitation input for subsequent vehicle system dynamic response analysis.



(**a**) (**b**)

Figure 3. Lateral CHSRTI spectrum and simulation verification. (**a**) Time-domain waveform of the lateral CHSRTI. (**b**) Comparison of the lateral CHSRTI spectrum with the theoretical analytical spectrum.

3. Method

This section details the proposed Physics-Constrained Dual-Attention Reinforcement Learning for Semi-Active Lateral Vibration Control of High-Speed Trains (SA-PRL). For-

mulated upon a quarter railway vehicle lateral model, this intelligent approach aims to effectively enhance vibration suppression performance.

3.1. Framework

The overall architecture of the proposed control method is depicted in Figure 4, comprising track excitation generation, vehicle dynamics simulation, and deep reinforcement learning, where solid arrows indicate information flow between components and different frame styles distinguish network modules from intermediate quantities. Track irregularity excitations derived from track spectra serve as inputs for the vehicle model to solve for dynamic responses. Based on vibration feedback, the control parameters are iteratively adjusted by the DRL algorithm under skyhook damping constraints to achieve the optimal control strategy.

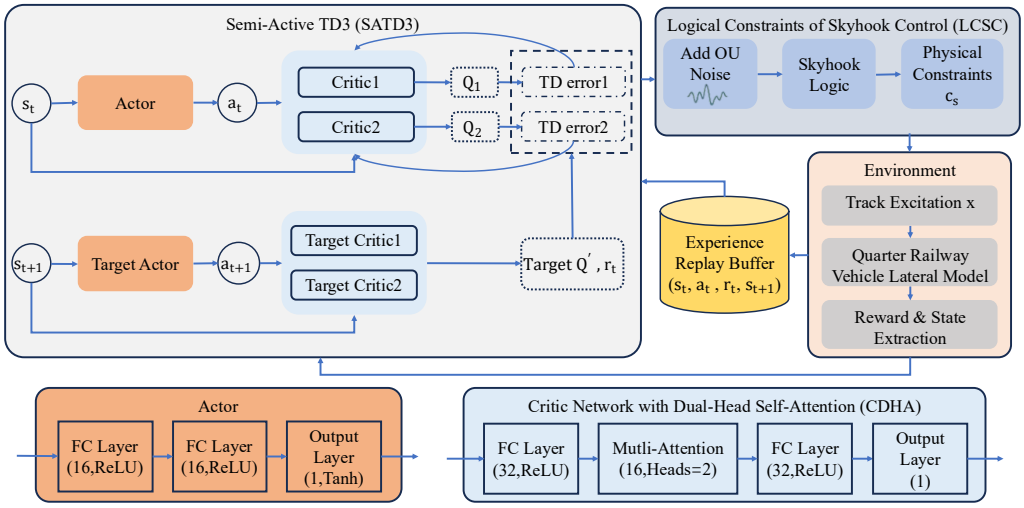


Figure 4. Overall architecture of the proposed SA-PRL framework.

This paper proposes SA-PRL, which utilizes the lateral vehicle dynamic model as its interaction environment. Built upon the Semi-Active TD3 (SATD3) algorithm, SA-PRL integrates both Logical Constraints of Skyhook Control (LCSC) and a Critic Network with Dual-Head Self-Attention (CDHA) to effectively enhance the lateral vibration suppression of the train.

During the training process, random sampling from an experience replay buffer is utilized to eliminate the temporal correlation of the data. Furthermore, a delayed policy update mechanism is implemented to ensure the sufficient convergence of the critic network prior to the iterative optimization of the policy network, significantly enhancing the optimization accuracy of the semi-active damping control strategy. External random irregularity excitations are generated based on track spectra. Under these excitations, the SA-PRL computes the lateral dynamic responses of the railway vehicle lateral mode in real time. High-dimensional state information incorporating dynamic features such as the displacements and velocities of the carbody and bogie at both current and historical time steps is obtained to construct the interaction environment between the reinforcement learning agent and the vehicle system. Based on the environmental state, the SATD3 outputs continuous and normalized control actions, thereby achieving the dynamic regulation of the train's secondary lateral damping.

Existing deep reinforcement learning methods directly output control actions without physical boundary constraints, which often leads to the generation of actions that violate the mechanical damping limits of semi-active suspensions. To address this, our proposed method processes the raw action output from the Actor network by first superimposing

Ornstein-Uhlenbeck (OU) random noise to facilitate exploration sampling. Subsequently, an LCSC mechanism is introduced to delineate the feasible damping adjustment range based on the relative motion patterns between the carbody and the bogie. The noise-augmented actions are then clipped and mapped, strictly confining the raw network output within the physically adjustable range of the semi-active damper. Finally, these actions are converted into damping coefficients that comply with the system's mechanical properties and can be realized in actual hardware. This process ensures that the generated control forces satisfy the engineering implementation requirements of the semi-active suspension system. Furthermore, to enhance the algorithm's perception and value-evaluation capabilities regarding vehicle dynamic features under complex operating conditions, this study integrates a dual-head self-attention mechanism into the Critic network, thereby constructing the CDHA.

3.2. Semi-Active Twin Delayed Deep Deterministic Policy Gradient

Targeting the vibration reduction mechanism of semi-active suspensions, this paper introduces modifications to the action output logic, policy update mechanism, and environment interaction paradigm of the existing TD3 algorithm. Consequently, the Semi-Active TD3 (SATD3) is proposed, achieving a deep integration between intelligent reinforcement learning optimization and physical semi-active constraints.

The overall framework of the proposed intelligent lateral vibration control algorithm for lateral vibration of the investigated railway vehicle model is illustrated in Figure 5. In this algorithmic framework, solid blue arrows indicate the direction of data flow, while dashed blue arrows denote the direction of network parameter updates. Built upon the Actor–Critic framework, the SATD3 algorithm incorporates a dual Critic network architecture to mitigate value overestimation, making it well-suited for the continuous regulatory control of semi-active damping. The Actor network takes the historical dynamic states of the vehicle as input and, following feature extraction, outputs a one-dimensional normalized continuous damping action. The dual Critic network concatenates the state and action as its input, embedding multi-head self-attention modules within its branches. Furthermore, the algorithm introduces mechanisms such as delayed policy updates and soft target network updates to enhance training stability under nonlinear track excitation environments.

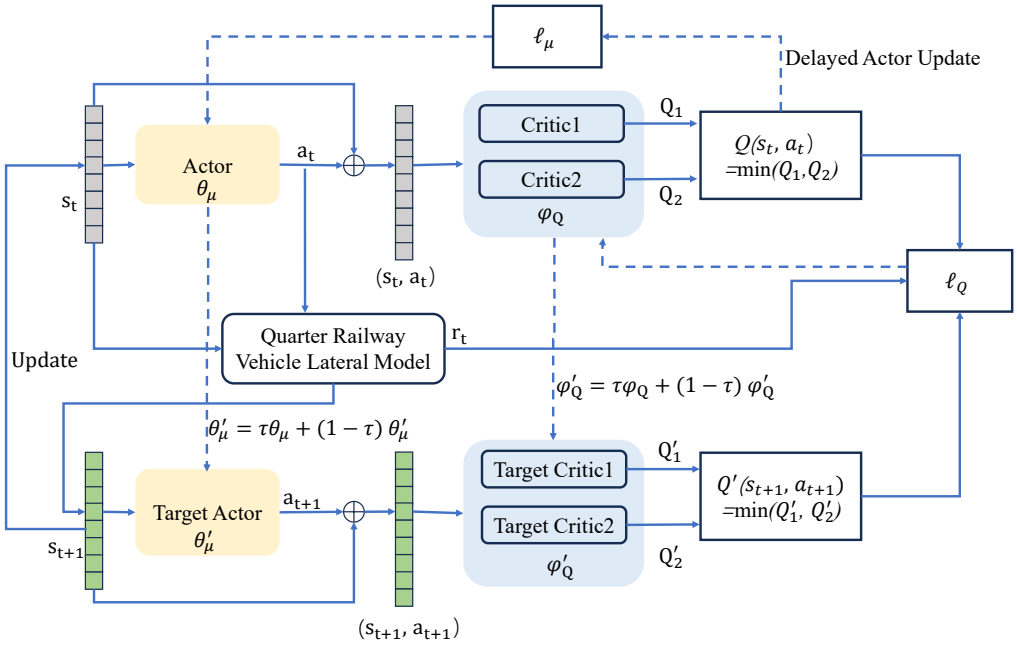


Figure 5. Training mechanism of the proposed SATD3 algorithm.

To guide the reinforcement learning agent in suppressing the lateral vibrations of the vehicle body, a reward function is designed with the objective of minimizing the lateral acceleration. The reward function is defined as $r_t = -\alpha|\ddot{x}_c|$, where $\ddot{x}_c$ denotes the lateral acceleration of the vehicle body at the current time step, and α is a weighting coefficient. By applying negative feedback to the acceleration magnitude, this reward function encourages the agent to learn a control policy that reduces vehicle vibrations.

To mitigate the value overestimation issue during training, SATD3 employs a double Critic network update mechanism. The target value function is constructed by taking the minimum output of the two target networks, which is formulated as:

$$y_t = r_t + \gamma \cdot \min\left(Q_1'(s_{t+1}, a_{t+1}), Q_2'(s_{t+1}, a_{t+1})\right) \tag{9}$$

where y_t is the target value at time step t; r_t is the single-step immediate reward; γ is the discount factor; and Q_1' and Q_2' are the outputs of the two target Critic networks. Taking the minimum of these outputs effectively suppresses the value overestimation phenomenon.

The parameters of the Critic network are updated by minimizing the mean squared error (MSE) between the predicted values of the Critic network and the target network, ensuring reliable value estimation. Concurrently, to address the susceptibility of semi-active damping outputs to gradient oscillations, this study introduces an L2 regularization constraint into the policy network's loss function. This imposes a smoothing constraint on continuous damping adjustment actions, preventing sudden damping variations and policy jumps during training, thereby enhancing the continuity and engineering practicality of the control output. The Critic loss function is defined as follows:

$$\ell_Q = \mathbb{E}[(Q(s_t, a_t) - (r_t + \gamma \cdot \min\left(Q_1'(s_{t+1}, a_{t+1}), Q_2'(s_{t+1}, a_{t+1})\right)))^2] \tag{10}$$

where $Q(s_t, a_t) = \min(Q_1, Q_2)$ is the value prediction of the current state–action pair by the Critic network; r_t is the immediate reward at the current time step; $Q'(s_{t+1}, a_{t+1})$ is the value estimation of the next state–action pair by the target Critic network; and γ is the discount factor. By minimizing the MSE between the predicted value and the target value, this loss function drives the Critic network to learn an accurate state–action value function.

To optimize the Actor network parameters and acquire a control policy, the deterministic policy gradient method is employed to update the Actor network. Simultaneously, an L2 regularization term is introduced to constrain variations in action amplitude, thereby enhancing the smoothness of damping adjustments. The policy loss function is formulated as:

$$\ell_\mu = -\mathbb{E}\left[Q_{\phi 1}\left(s, \theta_\mu(s)\right)\right] + \lambda \mathbb{E}\left[\theta_\mu(s)^2\right] \tag{11}$$

where $\mathbb{E}\left[Q_{\phi 1}\left(s, \theta_\mu(s)\right)\right]$ represents the expected state–action value corresponding to the action output by the policy network, serving as the core basis for policy optimization to achieve optimal action policy learning; $\mathbb{E}\left[\theta_\mu(s)^2\right]$ is the expected L2 regularization term for the output actions; and λ is the regularization weight coefficient.

This study implements a delayed policy update mechanism, wherein the policy network is updated only at fixed iteration intervals, while the Critic network is optimized continuously. This approach mitigates the training fluctuations caused by frequent policy variations. A soft update method is adopted for the target networks to synchronize the parameters. Specifically, the update formula for the target Actor network is:

$$\theta_\mu' = \tau\theta_\mu + (1 - \tau)\theta_\mu' \tag{12}$$

where θ_μ denotes the Actor network parameters; θ_μ' represents the target Actor network parameters; and $\tau \in (0, 1)$ is the soft update coefficient. This allows the target Actor network

parameters to smoothly track the Actor network updates, avoiding drastic fluctuations in the target values and ensuring stable convergence during the training process. The update formula for the target Critic network is:

$$\varphi'_Q = \tau\varphi_Q + (1-\tau)\varphi'_Q \tag{13}$$

where φ_Q denotes the Critic network parameters, and φ'_Q represents the target Critic network parameters. In summary, the proposed SATD3 algorithm successfully adapts the reinforcement learning framework to vehicle vibration suppression control. Furthermore, it lays a solid foundation for future research to integrate physical system constraints and introduce attention features for enhanced optimization of policy networks.

3.3. Logical Constraints of Skyhook Control

While the SATD3 algorithm achieves adaptive iterative policy optimization based on dynamic states, purely data-driven reinforcement learning models often optimize network output actions solely for vibration minimization. By neglecting the mechanical working principles of semi-active suspensions, these models are highly susceptible to generating invalid control variables that contradict the energy-dissipation characteristics of damping. Under the multi-coupled and strongly nonlinear lateral vibration conditions of trains, unconstrained policies are prone to chaotic damping regulation, reverse energy consumption, and the accumulation of invalid actions. These issues not only degrade training efficiency but also yield numerous physically unrealizable control solutions, thereby severely restricting the practical application of intelligent control strategies.

To address this, by integrating the well-established switching criteria of classical skyhook damping control, this paper designs a Logical Constraints of Skyhook Control (LCSC) mechanism to impose physical and mechanical constraints on the action outputs of the SATD3 algorithm. Consequently, an LCSC logic constraint module is constructed based on the skyhook vibration mitigation mechanism. This module performs secondary piecewise mapping and amplitude clipping on the actions generated by SATD3, transforming the unconstrained normalized actions into actual damping values that adhere to hardware limits and conform to the energy-dissipation laws of semi-active systems.To characterize the dynamic evolution patterns of the train's lateral vibration and provide adequate feature support for physical constraints and intelligent decision-making, this study constructs a dynamic state space encompassing both current and historical time-step information. The state vector simultaneously captures the time-series information regarding the lateral displacements and lateral velocities of the vehicle body and the bogie. This comprehensively reflects the coupled vibration state and motion trends of the vehicle-bogie system, effectively compensating for the inability of a single-step state to represent vibration velocity gradients and dynamic deviations. The specific state space is defined as:

$$\boldsymbol{s}_t = \left[x_b(t),\ x_c(t),\ \dot{x}_b(t),\ \dot{x}_c(t),\ x_b(t-1),\ x_c(t-1),\ \dot{x}_b(t-1),\ \dot{x}_c(t-1)\right]^{\mathrm{T}} \tag{14}$$

where x_b and x_c denote the lateral displacements of the bogie and the vehicle body, respectively; $\dot{x}_b$ and $\dot{x}_c$ represent the lateral velocities of the bogie and the vehicle body, respectively; t indicates the current time step; and $t-1$ represents the previous time step. This state space design encompasses the complete information of the current vibration state while incorporating the dynamic variation features from historical steps. It provides the reinforcement learning agent with rich dynamical information inputs, thereby enhancing the algorithm's perceptual capacity for vibration evolution patterns and establishing a solid data foundation for subsequent physical constraints and intelligent decision-making. Distinct from traditional unconstrained reinforcement learning methods that directly map control forces, the proposed LCSC mechanism utilizes the skyhook damping physical

criterion as a prerequisite constraint. It performs a secondary screening and mapping of the actions output by the SATD3 network to regulate the secondary suspension damping coefficient in real time, guaranteeing that the final control output complies with the physical properties of the semi-active suspension.

According to the lateral semi-active vibration mitigation mechanism of high-speed trains, when the vehicle body velocity and the relative velocity satisfy $\dot{x}_c(\dot{x}_c - \dot{x}_b) > 0$, the system is in an exacerbating vibration phase, requiring the suspension to provide variable damping to dissipate vibration energy. Conversely, when the system is in an attenuating vibration or ineffective energy-dissipation regime, no additional damping regulation is needed; the suspension maintains minimum damping to ensure the agile restoration of the vehicle body. The classical lateral skyhook semi-active control rule for railway secondary suspension is governed by the relative motion between the carbody and the bogie. The damping coefficient is adjusted based on the following conditions:

$$c_{\text{sky}} = \begin{cases} c_{\max}, & \dot{x}_c(\dot{x}_c - \dot{x}_b) > 0 \\ c_{\min}, & \dot{x}_c(\dot{x}_c - \dot{x}_b) \leq 0 \end{cases} \tag{15}$$

where $c_{\min}$ is the lower limit of the controllable damping and $c_{\max}$ is the upper limit of the controllable damping. Based on this physical mechanism, an adaptive damping output mapping function with boundary constraints is formulated as follows:

$$c_{\text{s}} = \begin{cases} c_{\min} + 0.5(a_t + 1)(c_{\max} - c_{\min}), & \dot{x}_c(\dot{x}_c - \dot{x}_b) > 0 \\ c_{\min}, & \dot{x}_c(\dot{x}_c - \dot{x}_b) \leq 0 \end{cases} \tag{16}$$

where a_t is the continuous action output by the policy network at time step t. Unlike the conventional two-state Skyhook controller, which determines the damping coefficient through discrete switching between the minimum and maximum damping values, the proposed framework employs the continuous action a_t to regulate the damping coefficient within the allowable range. This continuous control formulation allows the damping adjustment to vary according to the instantaneous vibration states and avoids the inherent discrete switching operation of the traditional Skyhook strategy. This architecture enables the reinforcement learning policy to perform refined optimization within the effective operating range of skyhook damping, thereby inheriting the stability of traditional algorithms while possessing the adaptive capabilities of intelligent control. Specifically, the physical rules are integrated into LCSC to realize the physics constraints. Damping modulation is only enabled when the carbody vibration is amplified to comply with semi-active dampers' energy-dissipation nature and avoid unrealistic energy injection. In addition, all damping outputs are bounded to $[c_{\min}, c_{\max}]$ to match the hardware limits of real dampers.

Compared with standard skyhook damping control, the proposed LCSC constraint mechanism employs the skyhook physical logic as the constraint boundary, allowing the SATD3 algorithm to autonomously learn optimal damping values in response to varying track excitations and vibration intensities. This ensures both physical realizability and adaptability to different operating conditions. This mechanism effectively circumvents the flaws of purely data-driven algorithms that detach from dynamic principles, and it resolves the issues of rigid parameters and poor condition adaptability in existing semi-active control systems. It provides a stable, reliable, and practically implementable physical constraint foundation for intelligent damping regulation under complex track excitations, while also generating high-quality decision samples for subsequent feature optimization via attention mechanisms.

3.4. Critic Network with Dual-Head Self-Attention

The SATD3 algorithm, in conjunction with the LCSC physical constraint mechanism, enables the adaptive optimal output and physically feasible control of semi-active damping. However, the lateral vibration of high-speed trains typically exhibits strong multi-state coupling, high randomness in operating conditions, and significant time-varying vibration characteristics. Existing fully connected Critic networks possess limited feature perception capabilities for high-dimensional dynamic states. They fail to effectively distinguish the correlation weights among vehicle body vibration, bogie vibration, and track excitation disturbances, rendering them prone to neglecting key coupling features. Consequently, this leads to insufficient value estimation accuracy and restricts the robustness of control decisions under complex operating conditions.

To resolve this issue, this paper constructs a Critic Network with Dual-Head Self-Attention (CDHA). By employing a multi-branch feature weighting mechanism, this network enhances the fine-grained extraction capability for multi-dimensional coupled dynamic information, thereby improving the algorithm's state perception and value discrimination accuracy under complex track excitations.The overall architecture and computational pipeline of the CDHA are illustrated in Figure 6. Circles with the same colour represent neurons within one hidden layer and the ellipsis omits parallel neurons. The state–action fused features are first fed into a basic fully connected layer to accomplish shallow feature extraction and dimension alignment. Subsequently, the preprocessed unified features are passed into a linear projection module to generate mutually independent Query, Key, and Value vectors, constructing the feature space required for attention computation. This study adopts a dual-head parallel attention architecture, establishing two independent attention branches to concurrently mine deep features and achieve the parallel analysis of multi-scale information. The computational formula for a single attention head is as follows:

$$\text{Attention}(Q_i, K_i, V_i) = \text{softmax}\left(\frac{Q_i K_i^{\mathrm{T}}}{\sqrt{d_k}}\right) V_i \tag{17}$$

where Q_i, K_i, and V_i are the Query, Key, and Value vectors corresponding to the i-th attention head, with $i \in \{1, 2\}$; $\sqrt{d_k}$ is the dimension scaling factor utilized to prevent gradient saturation caused by excessively large dot product values; the softmax function converts the correlation degrees between vectors into normalized weights, which are then used to compute a weighted sum of the Value vectors, yielding the output features of the single attention head.

The features output by the two distinct attention heads are initially concatenated along the channel dimension to achieve the fusion of multi-dimensional dynamic information. Following this, the concatenated combined features are fed into a learnable linear transformation layer for feature mapping and dimension integration. They are subsequently processed by a posterior fully connected layer to complete the high-dimensional nonlinear mapping. Ultimately, the network outputs the state value information, which is utilized for the Critic network loss calculation and the Actor network gradient updates. By fusing deep features through adaptive feature weighting, the CDHA cooperates with LCSC and the policy iteration of SATD3. Through the integration of the feature perception capabilities of the CDHA, the comprehensive framework concurrently guarantees physical feasibility, adaptive performance, and robustness across diverse operating conditions, thereby effectively realizing intelligent lateral vibration control for lateral vibration of the investigated railway vehicle model.

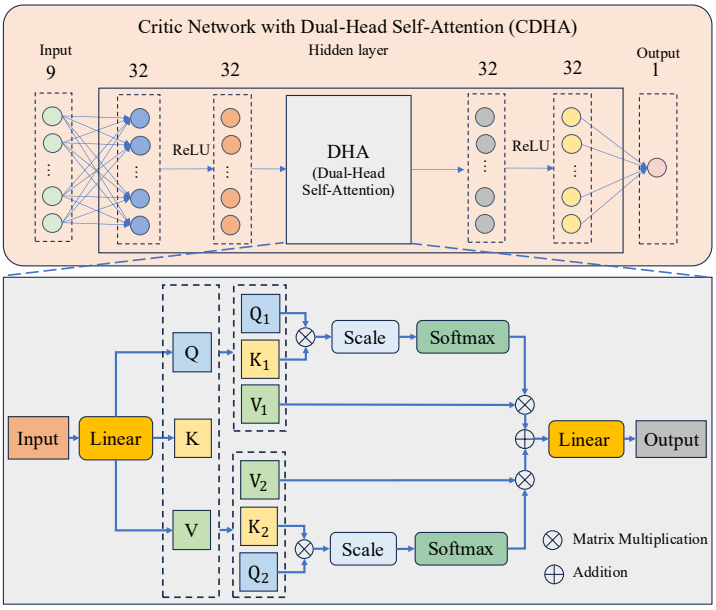


Figure 6. Architecture of the critic network with dual-head self-attention.

The pseudo code of the proposed SA-PRL method is given in Algorithm 1.

Algorithm 1 SA-PRL

Initialize actor network μ_θ, target actor $\mu_{\theta'}$ satisfying $\theta' \leftarrow \theta$; CDHA value Q_{ϕ_i}, target CDHA value Q'_{ϕ_i} satisfying $\phi'_i \leftarrow \phi_i$ $(i = 1, 2)$, experience replay buffer $\mathcal{D}$
Set hyper-parameters: discount factor γ, soft update rate τ, policy delay interval d
for each episode **do**
  Generate track excitation x_r, initialize system state s_t
  for each time step t **do**
    Select action with exploration noise: $a_t = \mu_\theta(s_t) + \mathcal{N}_t^{\mathrm{OU}}$
    Apply LCSC physical constraints to compute damping coefficient:
    $$c_s = \begin{cases} c_{\min} + 0.5(a_t + 1)(c_{\max} - c_{\min}), & \dot{x}_c(\dot{x}_c - \dot{x}_b) > 0 \\ c_{\min}, & \dot{x}_c(\dot{x}_c - \dot{x}_b) \le 0 \end{cases}$$
    Execute system dynamics: $s_{t+1} = f(s_t, c_s, x_r)$
    Obtain reward: $r_t = -\alpha|\ddot{x}_c|$
    Store transition (s_t, a_t, r_t, s_{t+1}) into replay buffer $\mathcal{D}$
    Sample mini-batch from $\mathcal{D}$
    Compute target action with smoothing noise: $a' = \mu_{\theta'}(s') + \epsilon$
    Compute target value: $y_t = r_t + \gamma \min_{i=1,2} Q'_{\phi_i}(s', a')$
    Update CDHA critic networks by minimizing critic loss ℓ_Q
    if t mod $d = 0$ **then**
      Update actor network by minimizing policy loss ℓ_μ
      Soft-update target networks: $\phi'_i \leftarrow \tau\phi_i + (1 - \tau)\phi'_i$
      $\theta' \leftarrow \tau\theta + (1 - \tau)\theta'$
    end if
  end for
end for

4. Simulation Verification and Analysis

4.1. Implementation Details

The simulation experiments comprise track spectrum generation, vehicle dynamic simulation, and control algorithm implementation. The simulation time step, train parameters, track excitations, and lateral dynamic model settings are kept consistent across

all simulation experiments to ensure a fair comparison. A dynamic random-seed mechanism based on training episodes is adopted, where the random seed is defined as the current training episode index. This strategy enables the generation of independent and diverse track irregularity excitations in each training episode. Meanwhile, the controlled randomization framework ensures overall reproducibility of the simulation experiments, thereby satisfying the requirement of experimental consistency. All reinforcement learning training and simulation experiments were performed on a workstation equipped with an Intel® Core™ i5-12490F processor, 32 GB RAM, and an NVIDIA GeForce RTX 2080 Ti GPU. The implementation was developed using Python 3.10 and PyTorch 1.12.1 under the Ubuntu 22.04 operating system with CUDA 12.8 support.

In all simulation experiments, the numerical integration time step is set to 0.005 s. The semi-active damper damping range is constrained within $c_{\min} = 1000\ \mathrm{N \cdot s/m}$ and $c_{\max} = 30{,}000\ \mathrm{N \cdot s/m}$ according to practical operating limitations. The passive damping coefficient is 30,000 $\mathrm{N \cdot s/m}$ [42]. Therefore, $c_{\max}$ is selected to provide comparable damping capacity to the passive suspension, while $c_{\min}$ is determined according to the minimum achievable damping level of practical semi-active dampers. The reinforcement learning action is normalized and clipped within $[-1, 1]$ and then mapped to the feasible damping range to ensure stable training and physically realizable control outputs.

After being trained on the lateral Chinese high-speed railway track irregularity (CHSRTI) spectrum, the proposed model is first evaluated under deterministic single-peak sinusoidal excitation and multi-peak sinusoidal excitation to investigate its vibration suppression mechanism at representative excitation frequencies. Subsequently, random track excitations are generated based on the lateral CHSRTI spectrum at 350 km/h and the lateral German low-interference track irregularity (GLITI) spectrum at 250 km/h. These tests are designed to assess the vibration reduction performance and generalization robustness of the proposed method under variations in both track spectrum characteristics and operating speeds. For all operating conditions, passive suspension, conventional Skyhook damping control, DDPG, and TD3 are selected for comparison.

To evaluate the vibration reduction performance of different control algorithms under various operating conditions, the root mean square (RMS) value of the carbody lateral acceleration is adopted as the primary performance metric. To mitigate the influence of randomness associated with a single track irregularity realization, the carbody lateral acceleration RMS is calculated using 10 independently generated track irregularity samples, and the average value is reported as the final result. This procedure improves the statistical reliability and generalization capability of the performance evaluation. To quantitatively assess the performance gain of each control algorithm relative to the passive suspension, the Improvement Ratio is employed as an auxiliary evaluation metric. Based on the RMS value of the carbody lateral acceleration, this metric characterizes the extent of vibration reduction achieved by a given control strategy compared with the passive suspension.

The network architecture and hyperparameters are summarized in Table 3. The Actor network consists of two fully connected hidden layers, each with 16 neurons. A piecewise-decaying exploration noise strategy is adopted during training. A relatively high noise level is applied during the first 50 training episodes to encourage sufficient exploration, after which the noise intensity is gradually reduced to facilitate stable policy convergence. The remaining hyperparameters, including the batch size, replay buffer capacity, learning rate, and discount factor, are systematically tuned to ensure both training efficiency and control performance under complex track excitation conditions. Furthermore, cosine annealing learning rate schedulers are employed for both the Actor and Critic optimizers. The learning rates are gradually reduced from their initial values to predefined minimum values over the training episodes, promoting stable optimization and improved convergence.

Table 3. Network architecture and hyperparameters of SA-PRL.

Category	Item	Value
Actor Network	Input dimension	8
	Output dimension	1
	Optimizer	Adam
	Initial learning rate	10^{-4}
Critic Network	Input dimension	9
	Output dimension	1
	Optimizer	Adam
	Initial learning rate	10^{-3}
Hyperparameters	Number of episodes	250
	Batch size	512
	Replay buffer size	10^5
	Discount factor (γ)	0.95
	Soft update rate (τ)	0.01
	Policy delay	2
	Exploration noise (1–50 episodes)	0.5
	Exploration noise (51–100 episodes)	0.3
	Exploration noise (101–250 episodes)	0.15

4.2. Results and Analysis

The cumulative reward curves obtained during training are shown in Figure 7, illustrating the convergence behavior of different control methods. As can be observed, the proposed method achieves a higher cumulative reward than the competing algorithms and exhibits stability during the later stages of training. These results indicate that the proposed method is capable of effectively learning an optimal damping regulation strategy under complex track excitation conditions.

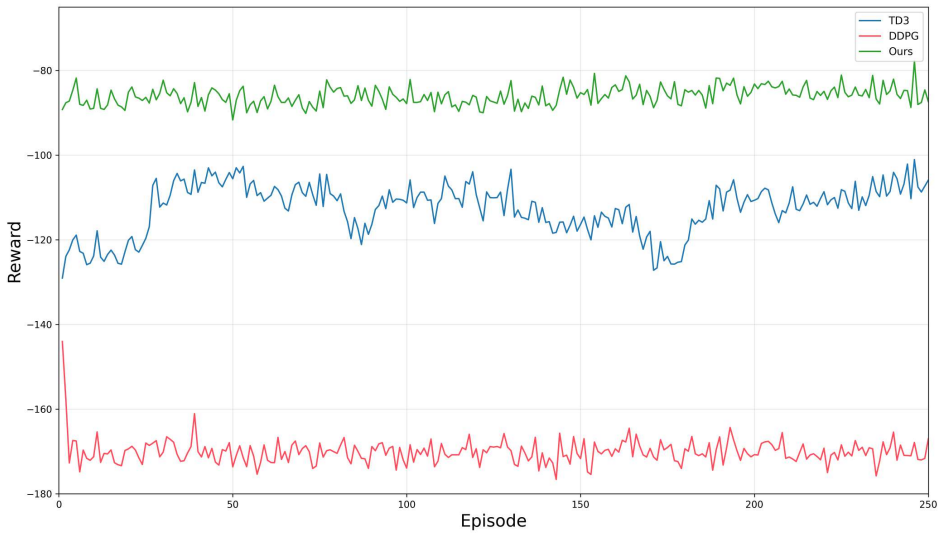


Figure 7. Comparison of cumulative reward convergence during training.

To evaluate the vibration reduction performance of the proposed method under transient impact excitations, two representative excitation scenarios, namely single-peak sinusoidal impact and multi-peak continuous impact, are constructed. The RMS values of carbody acceleration under different control strategies are summarized in Table 4.

Under the single-peak sinusoidal impact condition, all control strategies reduce the vibration level to varying degrees. Compared with the passive suspension, the RMS values of carbody acceleration under Skyhook, DDPG, and TD3 control decrease to 0.0615,

0.0650, and 0.0566, corresponding to improvement ratios of 15.9%, 11.1%, and 22.6%, respectively. In contrast, the proposed SA-PRL method achieves the best vibration reduction performance, reducing the RMS value to 0.0461 with an improvement ratio of 36.9%, outperforming all competing control strategies. Under the multi-peak continuous sinusoidal impact condition, the overall vibration levels increase because of the higher excitation intensity and longer excitation duration. The passive suspension yields an RMS value of 0.1911, while Skyhook control achieves a vibration reduction improvement of 16.5%. The reinforcement learning-based methods exhibit substantially effective performance, with the RMS values reduced to 0.1307 and 0.1303 under DDPG and TD3 control, respectively. The proposed SA-PRL method continues to deliver the best performance, further reducing the RMS value to 0.1170 and achieving an improvement ratio of 38.8%.

Table 4. Comparison of carbody lateral RMS acceleration (m/s^2) and improvement ratio under single-peak and multi-peak sinusoidal excitations.

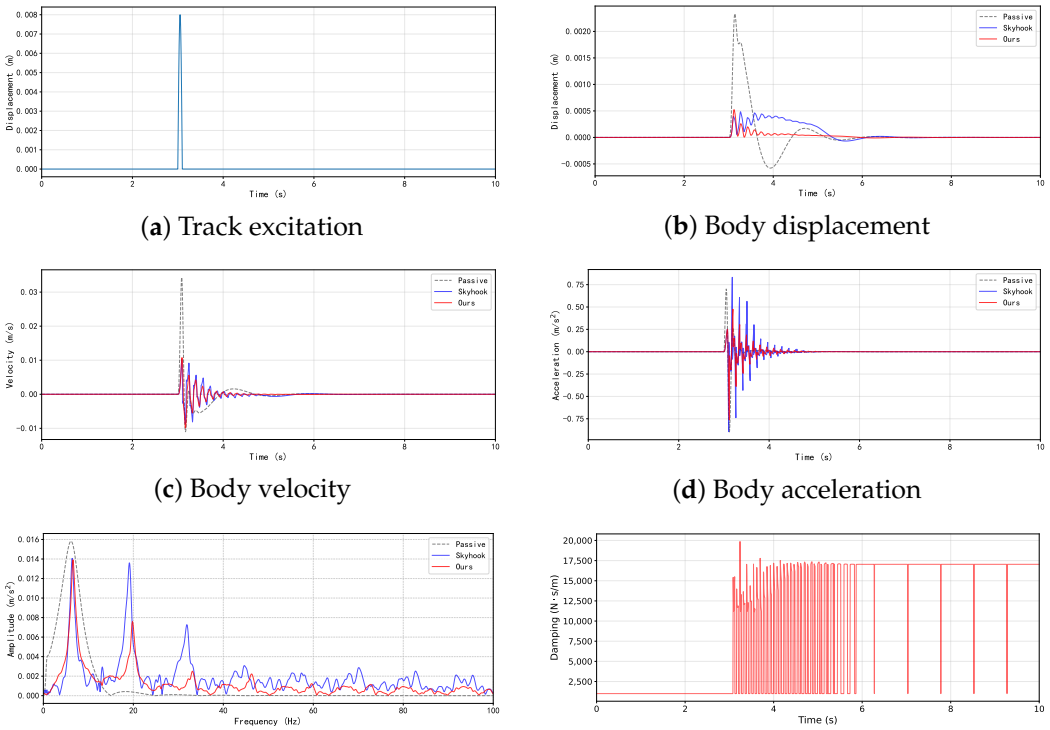
Type	Passive	Skyhook	DDPG	TD3	Ours
Single-peak Sinusoidal Excitation	0.0731	0.0615	0.0650	0.0566	**0.0461**
Multi-peak Sinusoidal Excitation	0.1911	0.1595	0.1307	0.1303	**0.1170**
Single-peak Improvement Ratio (%)	–	15.9	11.1	22.6	**36.9**
Multi-peak Improvement Ratio (%)	–	16.5	31.8	31.8	**38.8**

Note: Bold values indicate the best results.

Overall, compared with passive suspension, Skyhook control, the existing reinforcement learning algorithms DDPG and TD3, SA-PRL demonstrates vibration reduction capability under both single-peak and multi-peak impact excitation conditions.

Under the single-peak impact excitation condition, the proposed SA-PRL method is compared with passive suspension and Skyhook control. The track excitation and the corresponding time-domain responses distributions of the carbody acceleration are presented in Figure 8a–d. As shown in the figure, passive suspension and Skyhook control provide limited vibration attenuation, whereas the proposed SA-PRL method achieves the highest vibration reduction performance. From the time-domain perspective, under transient single-peak impact excitation, passive suspension and Skyhook control exhibit larger oscillation amplitudes and slower attenuation rates in both carbody velocity, displacement and acceleration responses. In contrast, SA-PRL rapidly suppresses the transient vibration peaks induced by the impact excitation, resulting in significantly reduced vibration amplitudes. From the frequency-domain perspective, Figure 8e shows that the amplitude spectrum of SA-PRL achieves lower spectral amplitudes than the Skyhook control strategy across most of the dominant frequency range, with particularly noticeable reductions around 20 Hz and 30 Hz, indicating enhanced vibration suppression.

Furthermore, the damping variation profiles shown in Figure 8f reveal that the proposed SA-PRL method can adaptively and continuously regulate the damping output according to the real-time vibration state of the carbody. Specifically, a large damping force is promptly applied when the impact occurs to suppress transient vibrations, while the damping level is reduced during steady-state operation to maintain ride comfort. This adaptive damping regulation capability is a key factor contributing to the vibration reduction performance of SA-PRL compared with the benchmark methods. Considering the time-domain responses, frequency-domain characteristics, and damping force outputs together, the results further verify the vibration attenuation capability of the proposed SA-PRL method under single-peak impact excitation.



(**a**) Track excitation (**b**) Body displacement

(**c**) Body velocity (**d**) Body acceleration

(**e**) Acceleration amplitude spectrum (**f**) Damping variation of our method

Figure 8. Dynamic responses under single-peak excitation.

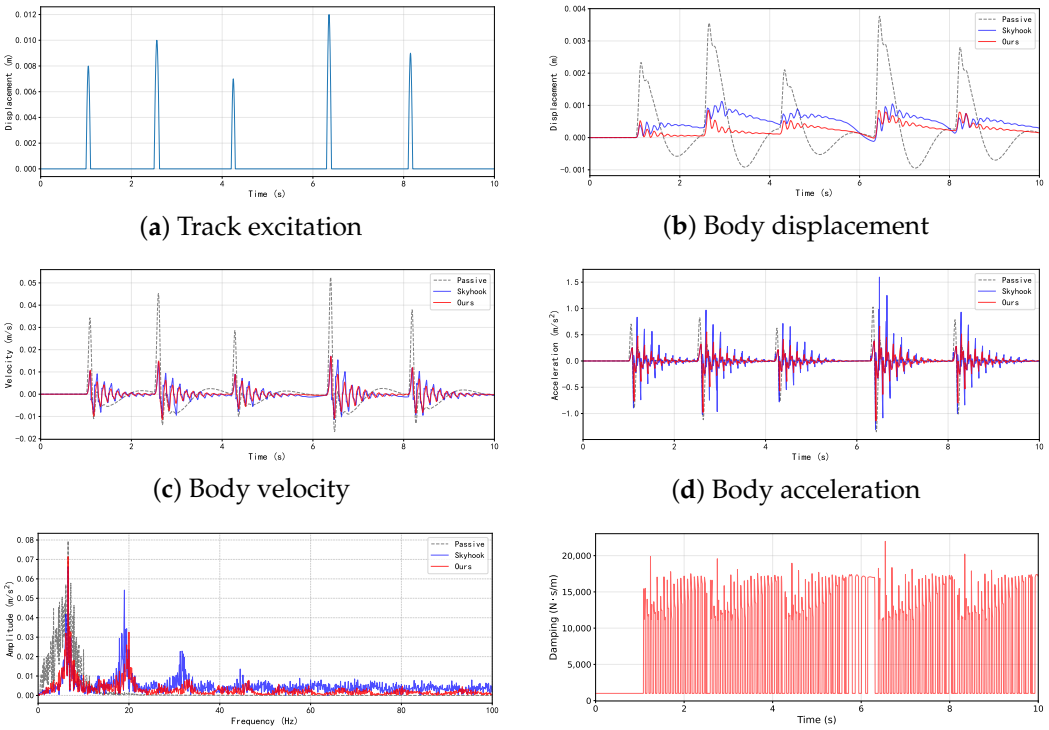
Under the multi-peak impact excitation condition, passive suspension, Skyhook control, and the proposed SA-PRL method are further compared. The track excitation and the corresponding time-domain responses distributions of the carbody acceleration are presented in Figure 9a–d. Consistent with the observations under the single-peak excitation condition, passive suspension and Skyhook control exhibit limited capability in suppressing continuous impact-induced vibrations, resulting in relatively slow attenuation of carbody oscillations. In contrast, the proposed SA-PRL method continues to demonstrate vibration reduction performance compared with Skyhook control. It effectively suppresses the vibration peaks induced by successive impact excitations in the time domain.

In addition, the frequency-domain characteristics presented in Figure 9e show that the proposed method achieves lower spectral amplitudes across most of the frequency range, especially around 20 Hz and 30 Hz, indicating enhanced vibration suppression. Benefiting from its flexible and adaptive damping regulation mechanism, as illustrated by the damping variation profile in Figure 9f, SA-PRL can continuously adjust the damping output according to the varying excitation characteristics, making it particularly outstanding for multi-peak alternating excitation conditions. As a result, the proposed method maintains stable vibration attenuation performance under continuous impact disturbances.

Overall, the results further confirm the effectiveness of SA-PRL in mitigating carbody lateral vibrations under complex multi-peak impact excitations, demonstrating its strong adaptability and robust vibration suppression capability.

Single-peak and multi-peak sinusoidal excitations are effective for evaluating the vibration suppression capability of control algorithms under representative fixed-frequency excitations and are therefore well suited for mechanism analysis. However, actual high-speed railway track excitations exhibit complex characteristics, including randomness, broadband frequency content, and multi-frequency coupling. Consequently, sinusoidal excitation alone cannot fully reproduce the complex vibration environment encountered during real train operation. To further approximate practical engineering conditions and comprehensively evaluate the vibration reduction performance and robustness of

the proposed SA-PRL method under realistic track disturbances, tests are conducted on 10 independently generated lateral CHSRTI samples.



(**a**) Track excitation (**b**) Body displacement

(**c**) Body velocity (**d**) Body acceleration

(**e**) Acceleration amplitude spectrum (**f**) Damping variation of our method

Figure 9. Dynamic responses under multi-peak excitation.

The corresponding results are summarized in Table 5. Passive suspension, Skyhook damping control, DDPG, and TD3 are selected as comparative methods. The RMS values of carbody lateral acceleration under different control strategies are statistically analyzed to quantitatively evaluate their vibration reduction performance. As can be observed from the results obtained under different track samples, the proposed method consistently achieves the lowest carbody acceleration RMS values across all random track excitation conditions, demonstrating effective stability and adaptability to varying operating conditions. The RMS value of the passive suspension remains approximately 0.1068 for different track samples, whereas Skyhook, DDPG, and TD3 control strategies all reduce the vibration level to some extent. Specifically, the mean RMS value is reduced to 0.0781 under Skyhook control, corresponding to an average improvement ratio of 26.8%. The DDPG method achieves a mean RMS value of 0.0777 with an improvement ratio of 27.8%, while TD3 further reduces the mean RMS value to 0.0635, yielding a vibration reduction improvement of 40.6% compared with the passive suspension. In contrast, the proposed SA-PRL method delivers the best performance under all test conditions. The mean RMS value of carbody acceleration is further reduced to 0.0561, corresponding to an improvement ratio of 47.5%. These results demonstrate that the proposed method not only effectively suppresses carbody lateral vibrations but also exhibits control stability and generalization capability under random track irregularity excitations.

Further analysis reveals that although Skyhook control can suppress vibrations to a certain extent through a fixed damping regulation strategy, its control logic and parameters remain relatively fixed, making it difficult to adapt to the dynamic variations of complex track excitations. Existing reinforcement learning algorithms possess a degree of adaptive optimization capability; however, they are not specifically designed to account for the physical constraints inherent to semi-active suspension systems, leaving room for further

performance improvement. In contrast, the proposed method integrates the constraint mechanism of Skyhook damping logic with the adaptive optimization capability of reinforcement learning, enabling precise adaptation to complex track excitation conditions and more effective vibration suppression. As a result, it consistently achieves vibration reduction performance and robustness across different track samples.

Table 5. Comparison of carbody lateral RMS acceleration (m/s^2) under CHSRTI excitations.

Method	Mean RMS Acceleration	Improvement Ratio (%)
Passive	0.1068 ± 0.0002	–
Skyhook	0.0781 ± 0.0021	26.8
DDPG	0.0777 ± 0.0066	27.8
TD3	0.0635 ± 0.0009	40.6
Ours	$\mathbf{0.0561} \pm 0.0013$	**47.5**

Note: Bold values indicate the best results.

To further illustrate the performance differences among the control strategies, boxplots of the carbody acceleration RMS values are constructed to visualize the distribution of vibration suppression performance under different track samples. The circles in the figure denote outliers, representing data points located beyond the interquartile range and the green dashed line denotes the mean value. The corresponding results are shown in Figure 10. It can be observed that, due to the absence of active damping regulation capability, the passive suspension exhibits a median carbody acceleration RMS significantly higher than those of the other control methods. Skyhook, DDPG, and TD3 control strategies achieve noticeable vibration reduction; however, their median RMS values remain higher than that of the proposed method. Among all compared approaches, the proposed SA-PRL method attains the lowest median acceleration RMS value, demonstrating a distinct advantage in vibration attenuation performance.

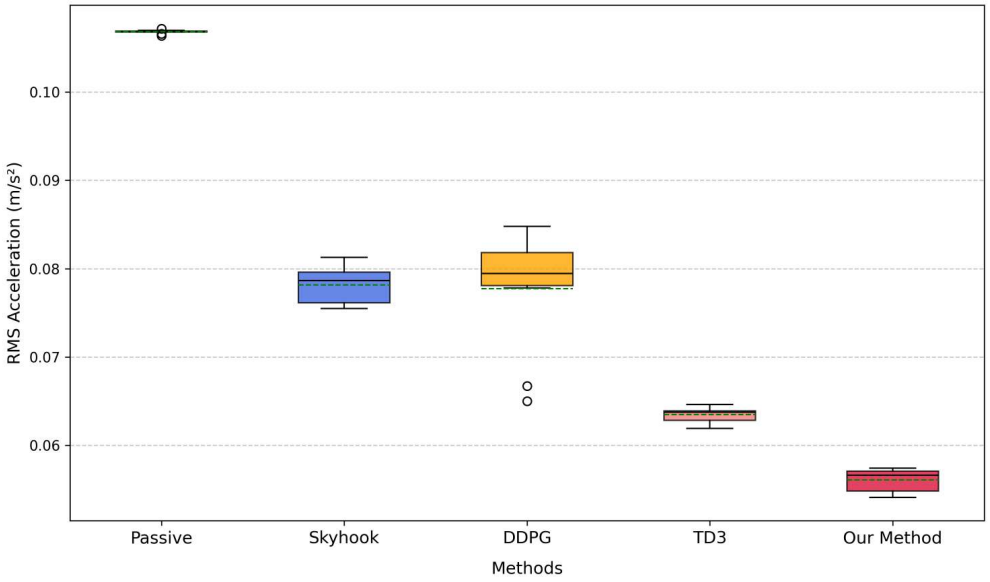
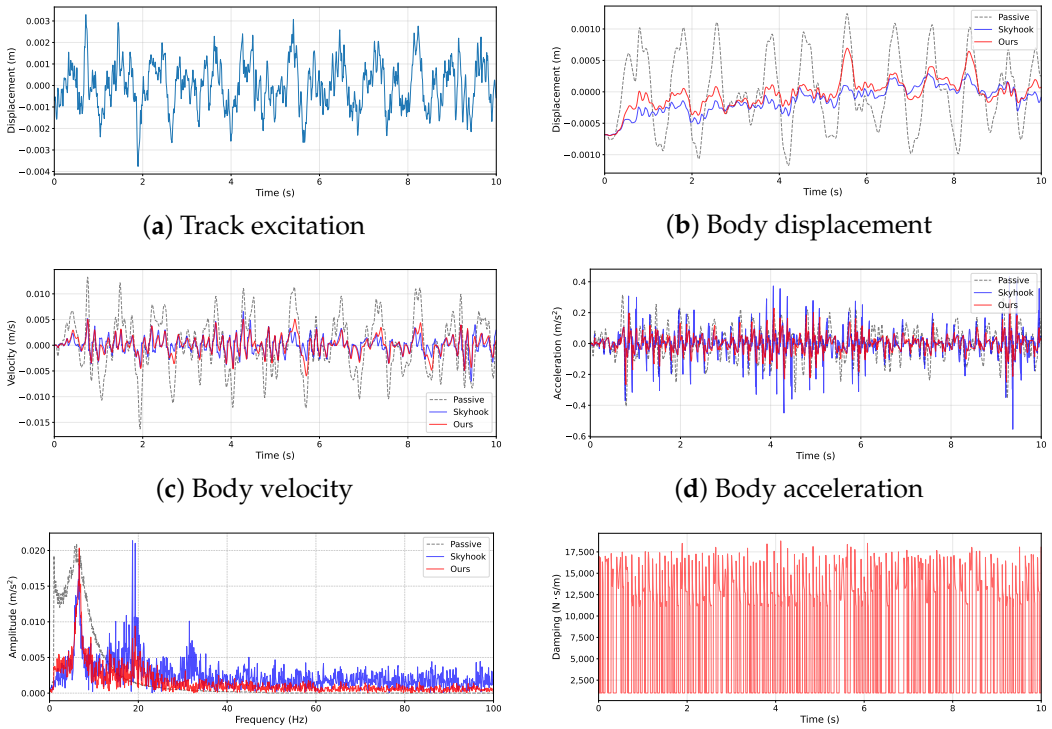


Figure 10. Comparison of body acceleration RMS distributions under 10 random lateral CHSRTI excitations.

To provide an intuitive comparison of the control performance, the time-domain responses of passive suspension, Skyhook control, and the proposed SA-PRL under the same track excitation are presented in Figure 11a–d.



(**a**) Track excitation (**b**) Body displacement

(**c**) Body velocity (**d**) Body acceleration

(**e**) Frequency-domain acceleration (**f**) Damping variation of our method

Figure 11. Dynamic responses under a random lateral CHSRTI excitation.

It can be clearly observed in Figure 11 that the passive suspension exhibits the most severe acceleration fluctuations, with the largest peak values and vibration amplitudes, resulting in the poorest vibration attenuation performance. In comparison, the proposed SA-PRL method produces the smoothest acceleration response, with vibration peaks significantly suppressed among all compared methods. The time-domain results demonstrate that SA-PRL achieves substantially vibration suppression performance than both passive suspension and Skyhook control. Compared with the Skyhook control strategy, the proposed method shows a slight loss in body displacement and velocity responses. However, it achieves a significant improvement in carbody acceleration, while the displacement and velocity responses remain much lower than those of the passive suspension. Therefore, SA-PRL exhibits the best overall vibration attenuation performance among the compared methods.

The frequency-domain characteristics presented in Figure 11e indicate that, compared with the Skyhook control strategy, SA-PRL achieves a pronounced reduction in the dominant spectral peak around 20 Hz, while exhibiting generally lower vibration amplitudes over the mid- and high-frequency ranges. Meanwhile, the damping coefficient variation in Figure 11f reflects the adaptive damping regulation behavior of SA-PRL, whereby the damping level is continuously adjusted within the allowable range according to real-time excitation conditions. These findings further verify the effectiveness of the proposed method in vibration reduction under lateral CHSRTI conditions.

To comprehensively assess the adaptive vibration suppression capability and engineering applicability of the proposed method under track irregularity excitations with different spectral characteristics, the lateral GLITI spectrum is introduced for cross-condition comparative testing. Specifically, tests are conducted on 10 independently generated GLITI samples. The corresponding RMS values of carbody acceleration are summarized in Table 6.

The results obtained from different random track samples show that all control strategies can reduce the vibration level of the carbody to varying degrees. However, the proposed method consistently achieves the lowest carbody acceleration RMS values across all

track samples, demonstrating effective stability and adaptability to different operating conditions. The passive suspension yields a mean RMS value of 0.2522, whereas the mean RMS values under Skyhook, DDPG, and TD3 control are reduced to 0.1799, 0.1687, and 0.1600, corresponding to the improvement ratios of 28.7%, 33.1%, and 36.6%, respectively. In comparison, the proposed SA-PRL method further reduces the mean RMS value to 0.1320, achieving an improvement ratio of 47.7% and delivering the best vibration reduction performance among all compared methods. These results indicate that the proposed method maintains effective vibration reduction performance when subjected to the investigated GLITI excitation.

Table 6. Comparison of carbody lateral RMS acceleration (m/s^2) under GLITI excitations.

Method	Mean RMS Acceleration	Improvement Ratio
Passive	0.2522 ± 0.0006	–
Skyhook	0.1799 ± 0.0037	28.7%
DDPG	0.1687 ± 0.0125	33.1%
TD3	0.1600 ± 0.0125	36.6%
Ours	$\mathbf{0.1320} \pm 0.0026$	**47.7%**

Note: Bold values indicate the best results.

To further illustrate the performance differences among the control strategies, boxplots of the carbody acceleration RMS values are constructed to visualize the distribution of vibration suppression performance under multiple operating conditions. The corresponding results are shown in Figure 12. It can be observed that, due to the absence of active damping regulation capability, the passive suspension exhibits a median carbody acceleration RMS significantly higher than those of the other control methods. Skyhook, DDPG, and TD3 control strategies achieve noticeable vibration reduction; however, their median RMS values remain higher than that of the proposed method. Among all compared approaches, SA-PRL attains the lowest median acceleration RMS value, demonstrating a clear advantage in vibration attenuation performance.

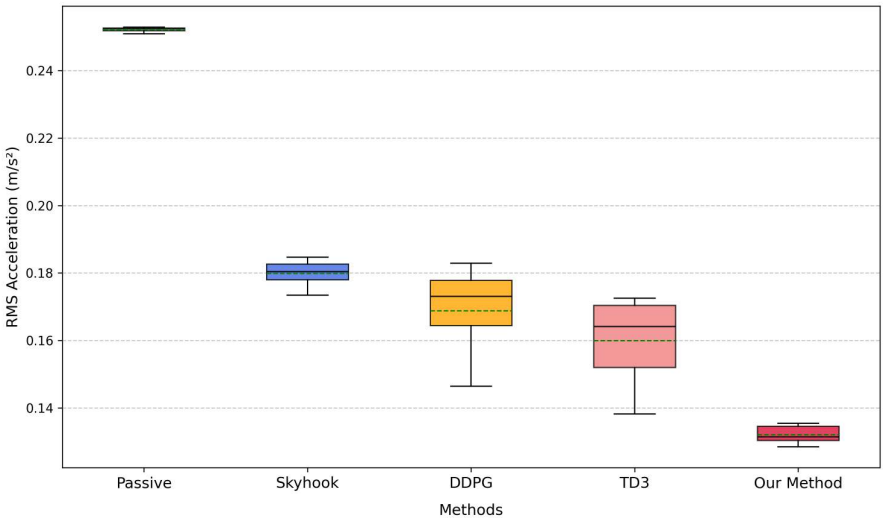
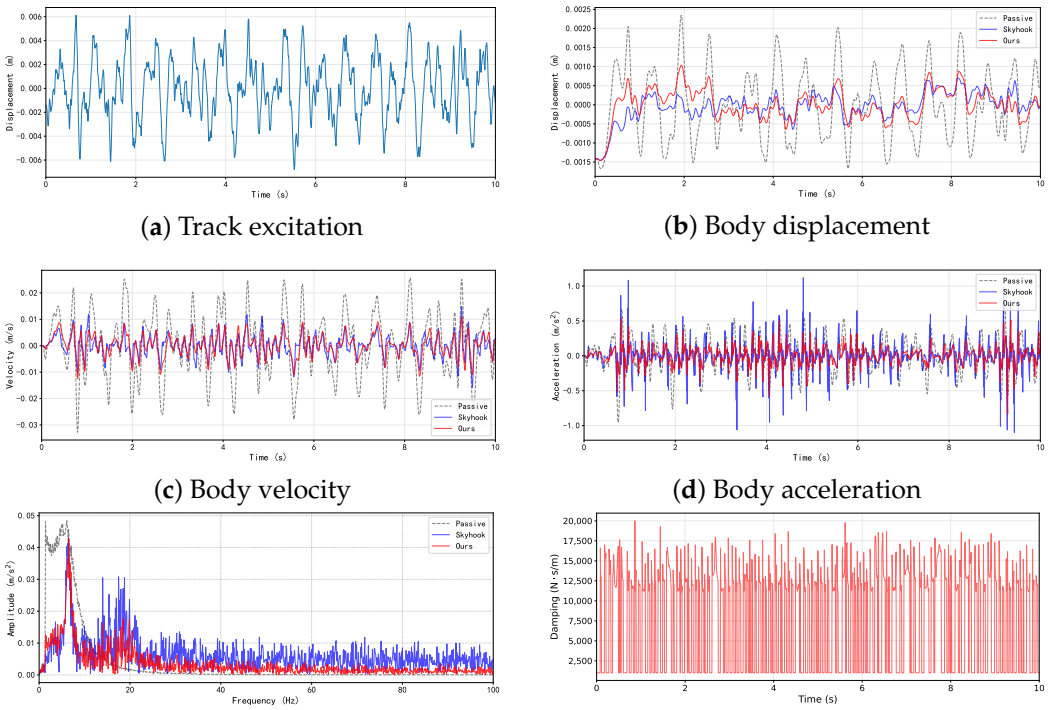


Figure 12. Comparison of body acceleration RMS distributions under 10 random lateral GLITI excitations.

Similarly, to provide an intuitive comparison of the control performance, dynamic responses of passive suspension, Skyhook control, and the proposed SA-PRL under the same track excitation are presented in Figure 13. The vibration reduction performance of passive suspension, Skyhook control, and the proposed SA-PRL method under the same track excitation is presented in Figure 13a–d. The passive suspension exhibits the most

severe acceleration fluctuations, with the largest peak values and vibration amplitudes, resulting in the poorest vibration attenuation performance. In contrast, the proposed SA-PRL method produces the smoothest acceleration response, significantly suppresses vibration peaks, and achieves the fastest response convergence among all compared methods. Its time-domain vibration suppression performance is better than that of both passive suspension and Skyhook control. Compared with the Skyhook control strategy, the proposed method exhibits a slight loss in body displacement and velocity responses, but achieves a much greater reduction in carbody acceleration, while both indices remain significantly lower than those of the passive suspension. Therefore, SA-PRL provides the best overall vibration attenuation performance among the compared methods.

The frequency-domain characteristics presented in Figure 13e show that, compared with the Skyhook control strategy, SA-PRL achieves lower spectral amplitudes over most of the frequency range, with particularly pronounced attenuation around 20 Hz. The damping variation profile in Figure 13f further reflects the adaptive damping regulation behavior of SA-PRL, which contributes to its effective suppression of carbody lateral vibrations and robust vibration attenuation performance under lateral GLITI excitation conditions. Overall, the proposed method demonstrates outstanding vibration attenuation capability by effectively reducing both the peak and overall levels of carbody acceleration, thereby contributing to improved ride comfort and operational safety.



(**a**) Track excitation (**b**) Body displacement

(**c**) Body velocity (**d**) Body acceleration

(**e**) Frequency-domain acceleration (**f**) Damping variation of our method

Figure 13. Dynamic responses under a representative GLITI excitation.

4.3. Ablation Studies

To separately investigate the individual contributions of the Semi-Active Twin Delayed Deep Deterministic Policy Gradient algorithm (SATD3), the Logic-Constrained Skyhook Control (LCSC) mechanism, and the Critic Network with Dual-Head Self-Attention (CDHA) module designed for semi-active damping regulation, a series of ablation experiments involving different module combinations is conducted. The results are presented in Table 7, where the mean RMS value of carbody acceleration over 10 randomly generated track irregularity samples is adopted as the evaluation metric.

Table 7. Ablation study of core modules in the proposed SA-PRL framework.

LCSC	SATD3	CDHA	Mean RMS Acceleration (m/s^2)	Improvement Ratio (%)
–	–	–	0.1068 ± 0.0002	–
✓	–	–	0.0781 ± 0.0021	26.8
–	✓	–	0.0635 ± 0.0009	40.6
✓	✓	–	0.0581 ± 0.0013	45.6
–	✓	✓	0.0601 ± 0.0017	43.7
✓	✓	✓	**0.0561** ± 0.0013	**47.5**

Note: "✓" indicates the inclusion of the component, "–" denotes absence and bold values indicate the best results.

Under the passive suspension condition, the average carbody acceleration RMS is 0.1068, yielding the poorest vibration reduction performance. When only the LCSC mechanism is employed, the mean RMS is reduced to 0.0781, corresponding to a vibration reduction improvement of 26.8%. Using only the SATD3 algorithm further decreases the mean RMS to 0.0635, achieving an improvement ratio of 40.6%, which outperforms single LCSC configuration. By incorporating LCSC into SATD3, the mean RMS is further reduced to 0.0581, corresponding to an improvement ratio of 45.6%. After additionally introducing the CDHA module, the system achieves the best vibration reduction performance, with the mean RMS reduced to 0.0561 and the highest improvement ratio of 47.5% relative to the passive suspension. According to the vibration evaluation criteria specified in ISO 2631 [44], the obtained RMS lateral acceleration corresponds to the not uncomfortable level.

To intuitively compare the performance differences across all ablation modules, Figure 14 visually illustrates the performance comparison of different module combinations in SA-PRL. Colored vertical bars on the left axis represent the mean RMS acceleration of carbody vibration, where smaller values indicate better control performance. The black polyline with circular markers on the secondary right axis quantifies the relative improvement ratio with respect to the passive baseline. The curve exhibits an overall upward trend, revealing that the stepwise addition of core modules continuously reduces vibration acceleration and improves suppression efficiency. The complete SA-PRL scheme integrating LCSC, SATD3 and CDHA corresponds to the shortest bar and the peak point of the improvement curve, which intuitively demonstrates its optimal vibration mitigation capability.

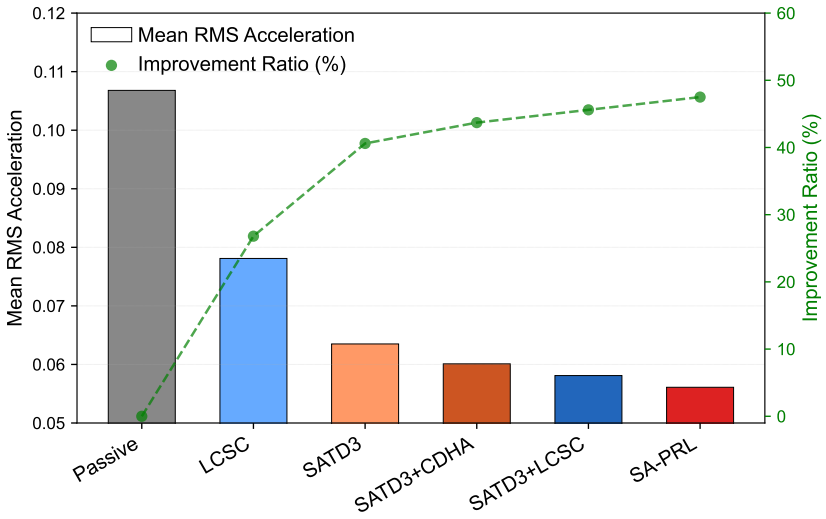


Figure 14. Performance comparison of different module combinations in SA-PRL.

Further analysis indicates that LCSC introduces skyhook damping logic constraints to provide a physically feasible mapping for the control output, thereby improving the vibration response to some extent. However, due to its relatively fixed control logic and

parameter settings, its adaptability to complex track excitations remains limited. In contrast, SATD3 fully exploits environmental state information to achieve adaptive optimization under complex track excitation conditions. Nevertheless, the absence of explicit physical constraints associated with semi-active suspension systems leaves room for further improvement in control performance. The integration of LCSC with SATD3 effectively enhances the stability and physical realizability of the control output, while the introduction of CDHA further strengthens the critic network's capability to perceive and evaluate complex dynamic state features. The synergistic effect of these three components ultimately enables the proposed SA-PRL method to achieve lateral vibration reduction performance under complex track excitation conditions.

We further investigate the effect of the reward weighting coefficient via simulation experiments. In all experiments with 10 randomly generated track irregularity samples, the LCSC mechanism and SATD3 algorithm are retained, and the corresponding results are presented in Table 8. The reward weighting coefficient α is found to have a noticeable influence on vibration reduction performance. When α is set to 0.1 , the improvement ratio reaches 43.3%. Increasing α to 1 further improves the performance, yielding the highest improvement ratio of 47.5%. However, when α is increased to 10, the improvement ratio decreases to 41.7%. Based on these results, α is set to 1 is selected as the optimal configuration for the proposed method.

Table 8. Effect of reward weight coefficient α on vibration reduction performance.

Weight Coefficient α	Mean RMS Acceleration (m/s^2)	Improvement Ratio (%)
–	0.1068 ± 0.0002	–
0.1	0.0605 ± 0.0013	43.3
1	$\mathbf{0.0561} \pm 0.0013$	**47.5**
10	0.0622 ± 0.0031	41.7

Note: Bold values indicate the best results.

To further investigate the influence of the suspension stroke constraint term in the reward function, an ablation study on the weighting coefficient β is conducted while fixing $\alpha = 1$, which was identified as the optimal value in the previous experiment. The reward function is defined as $r = -\alpha|\ddot{x}_c| - \beta|x_c - x_b|$, where $\ddot{x}_c$ denotes the carbody lateral acceleration and $|x_c - x_b|$ denotes the suspension stroke. The results are summarized in Table 9. When $\beta = 0$, the proposed controller achieves the lowest mean RMS lateral acceleration of 0.0561 m/s^2. As β increases, the RMS lateral acceleration gradually increases, indicating that introducing the suspension stroke penalty slightly compromises the vibration suppression performance by introducing an additional optimization objective. Meanwhile, the maximum suspension stroke decreases slightly from 6.44 mm at $\beta = 0$ to 6.30 mm at $\beta = 10$, demonstrating that a larger weighting coefficient can provide better control of suspension motion. Overall, the incorporation of the suspension stroke term leads to a trade-off between vibration reduction and suspension stroke regulation. Since the primary objective of this study is to minimize carbody lateral vibration and the maximum suspension stroke remains well within the safety limit under all tested settings, the final reward function adopts $\alpha = 1$ and $\beta = 0$.

Table 9. Effect of weight coefficient β on vibration reduction performance.

Weight Coefficient β	Mean RMS Acceleration (m/s^2)	Max Suspension Stroke (mm)
0	$\mathbf{0.0561} \pm 0.0013$	6.44 ± 0.0009
0.1	0.0580 ± 0.0014	6.43 ± 0.0008
1	0.0602 ± 0.0013	6.33 ± 0.0009
10	0.0602 ± 0.0013	$\mathbf{6.30} \pm 0.0007$

Note: Bold values indicate the best results.

To explore the influence of the number of attention heads in the CDHA module on control performance, a series of ablation studies with different attention head configurations is conducted. In all ablations, the LCSC physical constraint mechanism and the SATD3 algorithm are retained. The RMS value of carbody acceleration and the corresponding performance improvement relative to the passive suspension are used as evaluation metrics. The results obtained from the 10 random track irregularity samples are summarized in Table 10.

Table 10. Performance comparison of different numbers of attention heads.

Number of Heads	Mean RMS Acceleration (m/s^2)	Improvement Ratio (%)
–	0.1068 ± 0.0002	–
0	0.0601 ± 0.0013	43.7
1	0.0583 ± 0.0016	45.4
2	$\mathbf{0.0561} \pm 0.0013$	**47.5**
4	0.0584 ± 0.0012	45.3

Note: Bold values indicate the best results.

As shown in Table 10, the control performance first improves and then deteriorates as the number of attention heads increases. Without the attention mechanism, the mean RMS value of carbody acceleration is 0.0601 m/s^2, corresponding to an improvement ratio of 43.7% relative to the passive suspension. After introducing a single attention head, the RMS value decreases to 0.0583 m/s^2, and the improvement ratio increases to 45.4%. When the number of attention heads is increased to 2, the RMS value is further reduced to 0.0561 m/s^2, achieving the best performance with an improvement ratio of 47.5%. However, when the number of attention heads is further increased to 4, the RMS value rises to 0.0584 m/s^2, and the improvement ratio decreases to 45.3%.

These results indicate that an appropriate number of attention heads can effectively enhance the network's ability to capture and represent complex dynamic state features, thereby improving control performance. Nevertheless, an excessive number of attention heads increases model complexity and training difficulty, while also introducing a higher risk of overfitting, which may adversely affect policy learning and control effectiveness. Therefore, under the current experimental setting, the dual-head attention configuration provides the most favorable balance between feature representation capability and control performance.

To further investigate the influence of random initialization on the training stability of the proposed reinforcement learning controller, experiments with 10 different random seeds were conducted under the CHSRTI excitation. The RMS values of the lateral carbody acceleration are summarized in Table 11. It can be observed that the RMS acceleration values obtained with different random seeds remain within a relatively narrow range, varying from 0.0541 to 0.0574 m/s^2. The mean RMS acceleration is 0.0561 m/s^2 with a standard deviation of 0.0013 m/s^2, indicating limited performance fluctuation among different training runs. These results demonstrate that the proposed method maintains stable vibration reduction performance and that the reported results are not dependent on a specific random seed.

Overall, the ablation studies validate the effectiveness of the proposed core modules. In particular, the integration of the dual-head attention mechanism and the physics-constrained strategy further enhances control performance, enabling the proposed SA-PRL method to achieve superior vibration reduction capability and stronger generalization performance under complex track excitation conditions.

Table 11. Comparison of carbody lateral RMS acceleration (m/s^2) and improvement ratio under CHSRTI excitation for different random seeds.

Random Seed	Passive	Ours	Improvement Ratio (%)
seed = 1	0.1068	0.0555	48.0
seed = 2	0.1066	0.0545	48.9
seed = 3	0.1068	0.0541	49.3
seed = 4	0.1069	0.0546	48.9
seed = 5	0.1068	0.0574	46.3
seed = 6	0.1070	0.0569	46.8
seed = 7	0.1072	0.0574	46.5
seed = 8	0.1069	0.0570	46.7
seed = 9	0.1069	0.0571	46.6
seed = 10	0.1064	0.0563	47.1
mean $\pm$ std	0.1068 ± 0.0002	0.0561 ± 0.0013	47.5

5. Conclusions

To suppress carbody lateral vibrations during high-speed train operation, Physics-Constrained Dual-Attention Reinforcement Learning for Semi-Active Lateral Vibration Control of High-Speed Trains (SA-PRL) is proposed. Built upon the semi-active TD3 algorithm (SATD3), the proposed method incorporates Logical Constraints of Skyhook Control (LCSC) to construct a physics-constrained damping output mechanism, ensuring that the control actions always satisfy the mechanical characteristics and feasible damping boundaries of semi-active suspensions. In addition, a dual-head self-attention mechanism is embedded into the critic network to enhance the perception and value evaluation of the coupled dynamic interactions among the carbody, bogie, and track excitations, thereby improving control stability and policy optimization capability under complex operating conditions. Simulation experiments conducted under single-peak sinusoidal excitation, multi-peak sinusoidal excitation, Chinese high-speed railway track spectra and German low-interference track excitations demonstrate that the proposed SA-PRL method consistently achieves the best lateral vibration reduction performance compared with passive suspension, skyhook damping control, and conventional reinforcement learning approaches. Furthermore, ablation studies verify the proposed LCSC, the SATD3, and the CDHA. The results indicate that SA-PRL can autonomously learn an optimal semi-active damping regulation strategy, enabling adaptive control of the secondary lateral suspension damping under diverse track excitation conditions and effectively suppressing carbody lateral vibrations.

The proposed method provides a promising new framework for intelligent vibration control of high-speed train semi-active suspension systems, combining robustness and adaptability for engineering applications. Future work will extend the proposed SA-PRL framework to a lateral dynamic model of the full vehicle to account for more realistic vehicle–track interactions and coupled dynamic behaviors. This extension will enable a more comprehensive evaluation of the proposed controller under practical operating conditions and further verify its robustness and engineering applicability.

Author Contributions: Conceptualization, D.F. and L.G.; methodology, D.F. and L.G.; software, L.G.; validation, R.T., Y.Z. and Z.X.; formal analysis, D.F.; investigation, L.G.; resources, R.T.; data curation, L.G.; writing—original draft preparation, D.F. and L.G.; writing—review and editing, R.T.; visualization, L.G.; supervision, R.T.; project administration, R.T.; funding acquisition, Y.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (12402022) and the Science Research Project of Hebei Education Department (BJ2026297).

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: We sincerely appreciate the editors and reviewers for their valuable time and constructive comments.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

SA-PRL	Physics-Constrained Dual-Attention Reinforcement Learning for Semi-Active Lateral Vibration Control of High-Speed Trains
SA-TD3	Semi-Active Twin Delayed Deep Deterministic Policy Gradient
LCSC	Logical Constraints of Skyhook Control
CDHA	Critic Network with Dual-Head Self-Attention
DRL	Deep Reinforcement Learning
RMS	Root Mean Square
DDPG	Deep Deterministic Policy Gradient
TD3	Twin Delayed Deep Deterministic Policy Gradient
CHSRTI	Chinese High-Speed Railway Track Irregularity
GLITI	German Low-Interference Track Irregularity

References

1. Qi, Y.; Luo, S.; Zheng, Y.; Jing, H.; Wang, J.; Dai, H. Stability analysis of high speed trains based on active anti-yaw damper control with electro-hydraulic actuator. *Veh. Syst. Dyn.* **2026**, 1–27. [CrossRef]
2. Kang, W.; Zeng, J.; Mao, R.; Sang, H. Study on semi-active damping control strategies to improve dynamic behavior of high-speed trains subject to strong crosswinds. *J. Vib. Control* **2026**, *32*, 1597–1608. [CrossRef]
3. Zhang, H.; Ling, L.; Zhang, Z.; Zhai, W. Robust Constraint-Following Control for Bio-Inspired Structure Oriented Active Suspension System of High-Speed Trains. *IEEE Trans. Intell. Transp. Syst.* **2024**, *25*, 16107–16117. [CrossRef]
4. Jeniš, F.; Strecker, Z.; Michálek, T.; Macháček, O.; Staněk, P.; Malkovský, Z.; Žáček, J.; Kubík, M. Ride comfort improvement of a railway vehicle by semi-active damping control—Field tests. *Veh. Syst. Dyn.* **2025**, 1–18. [CrossRef]
5. Yang, S.; Zhao, Y.; Liu, Y.; Liao, Y.; Wang, P. A new semi-active control strategy on lateral suspension systems of high-speed trains and its application in HIL test rig. *Veh. Syst. Dyn.* **2023**, *61*, 1317–1344. [CrossRef]
6. Wang, Q.; Du, W.; Jiang, X.; Zeng, J.; Wei, L.; Peng, X. Operational Adaptability Improvement of High-Speed Trains Through Multi-Objective Control on Carbody Lateral Vibration. *Int. J. Struct. Stab. Dyn.* **2024**, *26*, 2650050. [CrossRef]
7. Shi, H.; Zeng, J.; Duan, L. Carbody swaying suppression for a high-speed rail vehicle by utilising active lateral suspension control. *Veh. Syst. Dyn.* **2025**, *63*, 1998–2017. [CrossRef]
8. Zhao, J.; Ling, L.; Zhang, H.; Zhang, Z.; Wang, K.; Zhai, W. Active Suspension and Linear Eddy Current Brake Control for Enhancing Ride Comfort and Safety of High-Speed Trains Equipped With Lift Airfoils. *IEEE Trans. Intell. Transp. Syst.* **2025**, *26*, 15028–15040. [CrossRef]
9. Shi, H.; Gan, F.; Zeng, J. Experimental Investigations of Valve-Controlled Dampers for Enhanced Ride Comfort on High-Speed Rail Vehicles. *IEEE Trans. Ind. Electron.* **2025**, *72*, 9465–9474. [CrossRef]
10. Zhang, Y.; Chen, C. A new mixed skyhook-backstepping semi-active control approach for high-speed train magnetorheological suspension. *J. Braz. Soc. Mech. Sci. Eng.* **2025**, *47*, 344. [CrossRef]
11. Yao, J.; Wang, S.; Chen, A.; Tian, J.; Wang, Z.; Guo, L. Semi-active actuator-based hybrid model predictive control for vehicle tilting. *Proc. Inst. Mech. Eng. Part D J. Automob. Eng.* **2026**, *240*, 2886–2896. [CrossRef]
12. Hua, Y.; Li, J.Y.; Zhu, S. Self-Powered Actively Controlled Lateral Suspensions of High-Speed Trains Using Energy-Regenerative Electromagnetic Damper and Model Predictive Control. *Int. J. Struct. Stab. Dyn.* **2024**, *26*, 2650066. [CrossRef]
13. Ling, L.; Zhao, J.; Zhang, H.; Zhang, Z.; Wang, K.; Zhai, W. Adaptive robust constraint-following control method for improving carbody hunting stability of high-speed trains. *Veh. Syst. Dyn.* **2026**, *64*, 73–94. [CrossRef]
14. Wu, H.; Gong, N.; Yang, J.; Gong, L.; Li, W.; Sun, S. Investigation of a semi-active suspension system for high-speed trains based on magnetorheological isolator with negative stiffness characteristics. *Mech. Syst. Signal Process.* **2024**, *208*, 111085. [CrossRef]
15. Kang, W.; Zeng, J.; Wang, Q.; Mao, R.; Yang, Y.; Sang, H. Dynamic performance improvement for high-speed trains under strong crosswind and tunnel passage using semi-active dampers. *Veh. Syst. Dyn.* **2026**, 1–23. [CrossRef]

16. Nhu, A.N.; Le, N.A.; Li, S.; Truong, T.D. Physics-Guided Reinforcement Learning System for Realistic Vehicle Active Suspension Control. In Proceedings of the 2023 International Conference on Machine Learning and Applications (ICMLA), Jacksonville, FL, USA, 15–17 December 2023; pp. 422–429.
17. Beller, S. Artificial intelligence supported semi-active suspension system design for motorcycles. *Results Eng.* **2025**, *28*, 107551. [CrossRef]
18. Tao, W.; Chen, B.; Zhou, L.; Zheng, Z.; Wu, J.; Duan, M. Design of a magnetorheological (MR) suspension damper for an agricultural tractor seat based on an adaptive neuro-fuzzy inference system (ANFIS) and active disturbance rejection control (ADRC). *Mech. Sci.* **2025**, *16*, 113–124. [CrossRef]
19. Zhang, Y.; Yan, X.; Yang, W.; Zhou, Y. Hierarchical Reinforcement Learning Based Optimal Control for Model-Free Linear Systems. *Mathematics* **2026**, *14*, 895. [CrossRef]
20. Hou, Y.; Liao, X.; Chen, G.; Chen, Y. Co-Evolutionary NSGA-III with deep reinforcement learning for multi-objective distributed flexible job shop scheduling. *Comput. Ind. Eng.* **2025**, *203*, 110990. [CrossRef]
21. Liu, C.; Chen, Y.; Xu, X. Fatigue life prognosis of composite structures using a transferable deep reinforcement learning-based approach. *Compos. Struct.* **2025**, *353*, 118727. [CrossRef]
22. Tanveer, J.; Lee, S.W.; Rahmani, A.M.; Aurangzeb, K.; Alam, M.; Zare, G.; Alamdari, P.M.; Hosseinzadeh, M. PGA-DRL: Progressive graph attention-based deep reinforcement learning for recommender systems. *Inf. Fusion* **2025**, *121*, 103167. [CrossRef]
23. Sun, Y.; Xia, H.; Su, C.; Zhang, R.; Wang, J.; Jia, K. A multi-agent enhanced DDPG method for federated learning resource allocation in IoT. *Comput. Commun.* **2025**, *233*, 108066. [CrossRef]
24. Chen, X.; Zang, J.; Zhao, F. Human-Like Autonomous Driving Car-Following Behavior Learning Based on Adversarial Training and Uncertainty-Aware DDPG. *IEEE Robot. Autom. Lett.* **2026**, *11*, 1754–1761. [CrossRef]
25. Nazemi Jenabi, M.; Asharioun, H.; Pourgholi, M. 3D UAV path planning based on an improved TD3 deep reinforcement learning for data collection in an urban environment. *J. Netw. Comput. Appl.* **2025**, *244*, 104336. [CrossRef]
26. Su, C.; Li, Y.; Wang, H.; Xia, Y.; Li, J.; Wan, G.; Chen, Y. AUV path planning in complex 3D underwater environments based on an improved TD3 algorithm. *Ocean Eng.* **2026**, *345*, 123688. [CrossRef]
27. Fujimoto, S.; van Hoof, H.; Meger, D. Addressing Function Approximation Error in Actor-Critic Methods. In Proceedings of the 35th International Conference on Machine Learning, Stockholm, Sweden, 10–15 July 2018; Volume 80, pp. 1587–1596.
28. Jia, C.; Liu, W.; He, H.; Chau, K. Superior energy management for fuel cell vehicles guided by improved DDPG algorithm: Integrating driving intention speed prediction and health-aware control. *Appl. Energy* **2025**, *394*, 126195. [CrossRef]
29. Tan, Z.; Wen, G.; Pan, Z.; Yin, S.; Wu, X.; Tohti, G. Control of a nonlinear active suspension system based on deep reinforcement learning and expert demonstrations. *Proc. Inst. Mech. Eng. Part D J. Automob. Eng.* **2024**, *238*, 4093–4113. [CrossRef]
30. Xia, G.; Liu, Y.; Zhang, Y.; Tang, X.; Wei, H.; Khajepour, A. Balancing Tracking Accuracy and Stability in Distributed-Drive Off-Road Autonomous Vehicles: An Adaptive TD3 Framework. *Proc. IEEE Trans. Veh. Technol.* **2026**, *75*, 12375–12386.
31. Liu, Q.; Xiong, P.; Zhu, Q.; Xiao, W.; Li, G.; Hu, G. A cooperative lateral and vertical control strategy for autonomous vehicles based on multi-agent deep reinforcement learning. *Proc. Inst. Mech. Eng. Part D J. Automob. Eng.* **2026**, *240*, 665–680. [CrossRef]
32. Zhan, L.; Xu, X.; Guo, X.; Deng, M.; Zou, J.; Li, W.; Du, H.; Li, Z. Deep reinforcement learning-based smart vibration control for magnetorheological suspension considering nonlinear dynamics. *Smart Mater. Struct.* **2025**, *34*, 065012. [CrossRef]
33. Jiang, M.; Gao, X.; Zhu, C. Designing optimal controller for semi-active rotor dry friction damper based on frequency-domain states: An offline deep reinforcement learning approach. *Mech. Syst. Signal Process.* **2025**, *238*, 113266. [CrossRef]
34. Wang, C.; Tao, G.; Cui, X.; Yao, Q.; Zhou, X.; Guo, K. Mechanism-data-driven control strategy for active suspension systems: Integrating deep reinforcement learning with differential geometry to enhance vehicle ride comfort. *Adv. Eng. Inform.* **2025**, *65*, 103326. [CrossRef]
35. Yang, J.; Wang, S.; Bai, F.; Wei, M.; Sun, X.; Wang, Y. Prior-Guided Residual Reinforcement Learning for Active Suspension Control. *Machines* **2025**, *13*, 983. [CrossRef]
36. Hu, T.; Jiang, L.; Pan, L.; Chen, B.; Gong, N.; Yang, J.; Gong, X.; Sun, S. Development of a semi-active suspension using a compact magnetorheological damper with negative-stiffness components. *Mech. Syst. Signal Process.* **2025**, *223*, 111842. [CrossRef]
37. Li, T.; Li, Y.; Wei, L.; Zhang, J. Comparative study on lateral vibration characteristics of the head and tail cars of high-speed trains under unsteady aerodynamic loadings. *Veh. Syst. Dyn.* **2025**, *63*, 518–536. [CrossRef]
38. Banerjee, C.; Nguyen, K.; Fookes, C.; Raissi, M. A survey on physics informed reinforcement learning: Review and open problems. *Expert Syst. Appl.* **2025**, *287*, 128166. [CrossRef]
39. Yang, X.; Zeng, S.; Liu, C.; Liu, X.; Wang, Z.; Yang, Y. Optimal Design and Performance Evaluation of HMDV Inertial Suspension Based on Generalized Skyhook Control. *Int. J. Struct. Stab. Dyn.* **2025**, *25*, 2550069. [CrossRef]
40. Jin, T.; Liu, Z.; Sun, S.; Ren, Z.; Deng, L.; Yang, B.; Christie, M.D.; Li, W. Development and evaluation of a versatile semi-active suspension system for high-speed railway vehicles. *Mech. Syst. Signal Process.* **2020**, *135*, 106338. [CrossRef]

41. Dimauro, L.; Venturini, S.; Tota, A.; Galvagno, E.; Velardocchia, M. High-speed tracked vehicle model order reduction for static and dynamic simulations. *Def. Technol.* **2024**, *38*, 89–110. [CrossRef]
42. Zhao, Y.; Yang, S.; Liu, Y.; Liao, Y. Simulation Analysis of Semi-Active Hybrid Control of Train Suspension System. *J. Dyn. Control* **2020**, *18*, 38–43.
43. Zhai, W. Vehicle-track coupled dynamics. In *Forthcoming*; Springer: Berlin/Heidelberg, Germany, 2015.
44. *ISO 2631-1:1997*; Mechanical Vibration and Shock—Evaluation of Human Exposure to Whole-Body Vibration—Part 1: General Requirements. ISO: Geneva, Switzerland, 1997.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.